DATA PROCESSING METHOD

5   CROSS-REFERENCE TO RELATED APPLICATIONS

Not applicable.

BACKGROUND OF THE INVENTION

The present invention relates to a data processing method and, more

10   particularly, to a data processing method facilitating adaptive data compression.

Digital data is typically used in a format that optimizes usability but often

requires data files that are at least twice as large as actually needed to represent

the information. Data compression is a general term for a number of algorithms

and computer programs that have been developed to convert data to a format

15   optimized for compactness.   In some instances, data compression can reduce the

quantity of data necessary to represent information by orders of magnitude.

Reducing the quantity of data processed and stored by a data processing system

is important because data transmission and storage are significant elements in

the processing time and cost of data processing systems. In some applications,

20   such as digital video transmission, the quantity of uncompressed data is so great

that practical systems would be cost and performance prohibitive without

substantial data compression.

Data compression methods are generally classified as lossy or lossless.

Lossless compression methods exploit statistical redundancy related to

25   similarities, correlation, and predictability of data.   When a lossless compression

method is applied to a data file, the data after decompression will be identical to

that of the original file. Data consistency is required for many types of data such

as executable program code, word processing files, or tabulations of numbers.

On the other hand, other types data files, including data for images and other

acquired signals, can often be approximated when decompressed without an appreciable loss. For example, a digital imaging device may capture textural detail that is unobservable a human viewer when included in a video frame. If the changes made to these signals resemble a small amount of noise, the resulting

5   distortion of the signal may go unnoticed or, at least, be tolerable. Lossy compression techniques exploit subjective redundancy in the data to identify and discard redundant and perceptually irrelevant information from the data file. Lossy compression typically compresses data more efficiently than lossless compression methods.

10   Many data compression schemes utilize several data compression techniques in a single composite process. For example, the MPEG-2 (ISO/IEC 13818) video compression standard utilizes a number of processes to compress the data within a single image (intraframe compression), between frames (interframe compression), and within the resulting data stream. Transform coding

15   utilizing the discrete cosine transform (DCT) is used to analyze and model the frequency of the signal representing blocks of pixels in the image. Transform coding is followed by quantization in which the transform coefficients are "rounded off" discarding part of the information necessary to perfectly reconstruct the original block of pixels. Temporal redundancy in the sequence of video images is

20   exploited by predictive interframe coding. The displacement of the image content between a reference frame and a predicted frame is estimated and quantified. The reference frame image is then displaced according to the estimate and the image content difference between the predicted frame and a motion compensated version of the reference frame or residual content is determined. The quantity of

25   data necessary to describe a sequence of frames is substantially reduced by reconstructing the predicted frame from a motion compensated version of the reference frame and the residual. In addition, the data stream containing the compressed intraframe and interframe data is typically compressed with a lossless method such as Huffman coding and run length encoding where

-2-

repeating data is replaced with shorter coded sequences.

While lossy techniques generally provide greater compression efficiency, the resulting data distortion can, at times, be objectionable. For example, the loss of information during the quantization of transform coefficients can produce

5   several compression artifacts in the decompressed image. These artifacts include the visibility of blocks used in the transformation process and noise along sharp edges in decompressed images. In addition, many compression techniques are more efficient when compressing certain types of data. For example, DCT transform coding is best suited to natural images with relatively gradual transitions

10 between areas of relatively consistent color. On the other hand, many images comprise a composite of elements including textual, graphical, and natural elements. To reduce objectionable compression artifacts and improve compression efficiency, images are often pre-processed and post-processed. To improve the efficiency of the compression process, the image may be examined

15 before compression and a "guess" made about the nature of the visual element to which a pixel or group of pixels belongs so that the image may be separated into components of different types and an appropriate process selected for each type of component. After the image is decompressed, its pixels may be examined to identify and smooth the color transitions between adjacent pixels. While image

20 pre-processing and post-processing is useful in increasing compression efficiency and repairing compression artifacts, it is computationally expensive, time consuming, and subject to errors.

What is desired, therefore, is a data processing method that facilitates adaptable and computationally efficient data compression utilizing efficient

25 methods.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic illustration of decomposition of an image into a array of pixels.

-3-

FIG. 2 is a schematic illustration of a method data processing according the present invention as applied to exemplary image pixel data.

FIG. 3 is a schematic illustration of a method of adaptive data processing.

5   DETAILED DESCRIPTION OF THE INVENTION

Information is an expression of an interaction between tangible and intangible objects. Digital computers and other digital electronic devices utilize information that can have only certain, fixed values. A fixed value of digital information or data is expressed as an ordered sequence of symbols. The value

10   of each symbol or digit is a function of the value of the symbol and the place it occupies in the sequence. The place in the sequence of symbols represents a power of a base number for a numbering system. For example, in the decimal number "12" the least significant digit "2" represents the quantity (2) of units (10^0) and the most significant digit "1" represents the quantity of tens (10^1) included in

15   the value of the sequence of symbols. Digital electronic devices typically utilize binary numbers where a symbol in the sequence or bit represents one of two possible states or values (1 or 0) and the position of the bit in the sequence of bits indicates a power of the base number "2". For example, a sequence of eight bits or a byte can represent values between 0 and 255. Digital information or data

20   typically comprises a binary number expressed as ordered sequence or block of bits.

Lossless compression is a process that reduces the number of bits required to represent a block of data, but permits the bits of the original block to be restored exactly. Lossless compression exploits statistical redundancy in data

25   and reduces the quantity of data by identifying and coding data similarities, correlation, and predictability. For example, run-length encoding is a lossless compression process that is useful in compressing data that contains repeated occurrences of the same symbol or group of symbols. For instance, an image of the night sky would include long runs of the symbol or symbols that represent the

pixels of a black background. When a run of the appropriate symbol is detected, the encoder outputs the value of the symbol as a flag indicating the start of the run and follows the flag with the number of symbols included the run of symbols. Huffman encoding is another commonly used lossless compression technique. In

5   this technique unique variable length code symbols are assigned to patterns symbols in the data according the frequency of repetition of the pattern. For example, if the letter "e" is the most frequently occurring symbol in a text document its bit sequence would be assigned the shortest code symbol (for example a single bit "1"). Less frequently used sequences are assigned longer,

10 unique code symbol sequence. If the average length of the code sequence of the encoded data is shorter than the average length of the symbol sequences in the original data, the total amount of data will be reduced.

Lossy compression exploits subjective irrelevancy in the data. Lossy compression is typically applied to digitized signals, such as audio signals or

15 image pixel values, where a degradation of the decompressed signal, with respect to the original signal, is perceptually irrelevant or, at least, tolerable. For example, lossy compression is often used in image processing where inexactitude in the color of a pixel is expected or, at least, tolerable and where the quality of the reconstructed image can be traded for compression efficiency. Lossy

20 compression is typically a two-step process in which useful components of the data are identified and modeled and a portion of the data, considered to be perceptually irrelevant to the data's destination, is discarded. In image processing, adjacent pixels tend to be highly correlated. Transform coding is typically used to model image pixels and analyze the frequencies included in the

25 visual signal. Quantization of the transform coefficients compresses the data by discarding a portion of the information necessary to perfectly describe the pixels. While the pixels of the image cannot be precisely reconstructed, the distortion produced by the loss of information may not be perceptually objectionable to a human viewer.

Data is typically generated, processed, and stored as a sequential progression of data blocks each comprising an ordered sequence of symbols representing the value of the sequence. For example, to compress and store an image, the image is scanned and decomposed to an array of pixels. In the case

5   of a color image, the image is typically filtered and separated into a plurality (typically three or four) color planes. Each color plane is a gray scale image comprising pixels spatially corresponding to those of the scanned image. A pixel of a color plane has a luminous intensity or luminosity proportional to the intensity of the particular color component at that pixel. The image can be described by an

10   array of symbol sequences. Each sequence comprises a specific number of bits or symbols expressing a value corresponding to the intensity of a pixel in the appropriate color plane. For example, an image may be separated into red, green, and blue color planes and the intensity of each pixel in each color plane represented by eight bits. If this color data is interlaced, the image can then be

15   described by a data stream comprising a sequence of data blocks with each data block comprising an ordered sequence of 24 bits describing the intensity of the three color components making up the image. While compression algorithms typically detect and exploit irrelevancy and statistical redundancy in the stream of data, the present inventors discovered that compression efficiency can be often

20   be improved and computational requirements reduced by partitioning the data along bit planes comprising the symbols holding one or more places in the plurality of ordered sequences and processing the data comprising bit planes.

FIG. 1 schematically illustrates the application of the data processing method in an image processing application. An image 20 is decomposed into an

25   array of picture elements or pixels 22. The image 20 is typically a gray scale image. However, the gray scale image 20 may be one of the constituent color planes that are superimposed to create a color image. Color images are commonly displayed by superimposing red (R), green (G), and blue (B) color planes or by superimposing cyan (C), magenta (M), yellow (Y), and black (K) color

planes. Referring to FIG. 2, the luminous intensity of each pixel 22 of the image 20 is expressed as an ordered sequence of symbols 24 (indicated by a bracket) having four binary bits. A pixel having a four bit value can have an intensity equivalent to one of sixteen shades of gray ranging from black to white.

5 For convenience, the intensities of the pixels 24 are illustrated as a rectangular array 26 of a plurality of ordered symbol sequences or blocks even the blocks are typically processed a sequential series of blocks. The present inventors recognized that while there is often statistical redundancy in the symbols of the serial sequence of data blocks, such as the repetition of the pixel value for

10 the second 28 and third 30 pixels, there is often greater redundancy in the bits or symbols occupying a particular place or position in the plurality of ordered sequences. For example, the most significant bits (MSB) 32 of all the exemplary pixel values are the same. Likewise, the subjective effect of the value of a bit on the decompressed data is likely correlated to the bit's position in the ordered

15 sequence. For example, pixels representing the sky in an image will be predominantly blue. Local variations in the shade of the sky will generally be accounted for by intensity variations expressed by changes in the least significant bit (LSB) 34 or a plurality of bits of lessor significance. A change in the MSB 32 for the four bit pixel values would correspond to a 50% variation in intensity which

20 is unlikely for large areas of an image.

In the method of the present invention, the data of the array 24 is partitioned along bit planes containing the bits occupying a particular place or plurality of places in the ordered sequences making up the array. For example, the most significant bits of the bit sequences may be grouped in a single bit

25 plane 36. Likewise, the least significant bits or the two least significant bits might occupy a bit plane 38. For convenience, the bits included in the bit plane can be assigned to a file and processed. Lossless or lossy compression algorithms are applied to the data in the file to reduce the quantity of data in the file. An algorithm can be selected that is likely to produce high compression efficiency or

optimize compression efficiency and data distortion. For example, applying run length encoding to the bit plane 36 comprising the most significant bits of the data of the array 26 may achieve high efficiency because there are likely to be substantial runs of like bits. Lossless compression of the most significant bit

5   plane 36 avoids distortion of the decompressed data and the possibility of substantial changes in the data and, for example, the color of the pixel represented by the data. On the other hand, a different compression algorithm may be more appropriate to a bit plane comprising least significant bits, such as bit plane 38. In the case of image data, a lossy compression technique applied to

10   a bit plane of less significant bits may provide high compression efficiency without substantially subjectively impacting quality of the decompressed image.

Since the bit planes may be processed by different methods, a data processing device that includes a library of compression algorithms can apply a plurality of algorithms to each plane for adaptive processing. Referring to FIG. 3,

15   the symbols of a bit plane can be input 50 to a plurality or processing methods 52, 54, and 56. The results of the plurality of processes can be evaluated in a comparator 58 and the better result, either the smallest quantity of compressed data or the result optimizing another criterion, retained 60 and stored. Each bit plane may be processed sequentially in a single processor or

20   the bit planes may be simultaneously input to several processors for parallel compression by a plurality of methods as illustrated in FIG. 3.

Data is typically processed as a series of ordered sequences or blocks. While significant compression can be achieved, data often has greater statistical redundancy when partitioned along a bit plane including the symbols occupying a

25   particular place or places in the sequence of ordered symbols. This redundancy can be exploited to achieve high efficiency, adaptive compression and reduce computational complexity.

The detailed description, above, sets forth numerous specific details to provide a thorough understanding of the present invention. However, those

skilled in the art will appreciate that the present invention may be practiced without these specific details. In other instances, well known methods, procedures, components, and circuitry have not been described in detail to avoid obscuring the present invention.

5   All the references cited herein are incorporated by reference.

The terms and expressions that have been employed in the foregoing specification are used as terms of description and not of limitation, and there is no intention, in the use of such terms and expressions, of excluding equivalents of the features shown and described or portions thereof, it being recognized that the

10   scope of the invention is defined and limited only by the claims that follow.